

EXPRESS MAILING CERTIFICATE

"EXPRESS MAIL" Mailing Label No.: ELO84649384US

Date of Deposit : August 27, 2001

I hereby certify that this paper or fee is being deposited  
with the United States Postal Service "Express Mail Post Office to  
Addressee" service under 37 CFR 1.10 on the date indicated above  
and is addressed to the Assistant Commissioner for Patents, Box  
Patent Application, Washington, D.C. 20231

Typed or printed name of person signing this certificate:



**24024**

PATENT TRADEMARK OFFICE

Signed: Candice umore

**PATENT**

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

In re Application of	)	Examiner: Not yet assigned
Gorski, et al.	)	
	)	
Serial No.: Not yet assigned	)	Group Art Unit: Not yet assigned
Parent Serial Number: 09/078,465	)	
	)	
Filed: Concurrently herewith	)	
Parent Filed: May 14, 1998	)	
	)	
For: <b>GROWTH ARREST HOMEobox</b>	)	Attorney Docket No.: 22311/04015
<b>GENE</b>	)	

Assistant Commissioner for Patents  
Washington, D.C. 20231

**FIRST PRELIMINARY AMENDMENT**

Dear Sir:

Please amend the above-identified continuation application as follows:

In the Specification:

Page 1, before "Background of the Invention" insert

Cross-Reference to Related Applications

This application is a continuation of the co-pending, commonly assigned, United States Patent Application Serial No.: 09/078,465, filed on May 14, 1998, which is a continuation of United States Application Serial No. 08/203,532, filed on February 24, 1994, and which issued as U.S. Patent No. 5,856,121 on January 5, 1999.

Page 1, line 21, delete “extra cellular” and insert ---extracellular---.

Page 2, lines 23 to 24, delete “extra cellular” and insert ---extracellular---.

Page 3, line 14, delete “the”.

Page 3, line 15, delete “The”.

Page 4, line 6, after “sequence” add ---SEQ ID NO: 1---.

Page 4, line 7, after “sequence” add ---SEQ ID NO: 2---.

Page 4, line 20, after “sequence” add ---SEQ ID NO: 3---.

Page 4, line 21, after “sequence” add --- SEQ ID NO: 4 ---.

Page 7, line 16, delete “nucleotide” and insert ---nucleotide---.

Page 8, line 9, after “sequence” add --- SEQ ID NO: 5---.

Page 8, line 36, after “gene” add ---SEQ ID NO: 1---.

Page 9, line 9, after “protein” add --- SEQ ID NO: 2---.

Page 9, line 19, delete the word “phosphorylated”.

Page 12, line 20, after “sequence” add ---SEQ ID NO: 3---.

Page 15, lines 2 and 3, delete the word “follows”.

Page 16, line 6, after “P2B” add ---SEQ ID NO: 6---.

Page 16, line 7, after “H2” add ---SEQ ID NO: 7---.

Page 16, line 8, after “H2R” add ---SEQ ID NO: 8---

Page 16, line 9, after “H3” add ---SEQ ID NO: 9---.

Page 16, line 10, after “H6” add ---SEQ ID NO: 10---.

Page 16, line 11, after “H7” add ---SEQ ID NO: 11---.

Page 16, line 12, after “AP” add ---SEQ ID NO: 12---.

Page 16, line 13, after “Anchor” add ---SEQ ID NO: 13---.

Page 18, line 16, delete “bank” and insert ---band---. Id.

Page 22, line 26, after “3” insert ---SEQ ID NO: 14---.

Page 22, line 27, after “3” add ---SEQ ID NO: 15---.

Page 23, line 37, delete “6752,” and insert ---6762---.

Page 29, line 1, delete “The”.

page 31, line 16, delete “oligo sequence” and insert ---oligosequence---.

Page 31, line 19, after “3” add ---SEQ ID NO: 16---.

Page 31, line 20, after “3” add ---SEQ ID NO: 17---.

Page 31, line 21, after “3” add ---SEQ ID NO: 18---.

Page 31, line 22, after "3" add ---SEQ ID NO: 19---.

Page 33, please replace the Sequence Listing with the Sequence Listing which is attached hereto.

IN THE CLAIMS

Please amend the claims as follows:

1. (Amended) An isolated DNA encoding a mammalian Gax protein [or portion thereof] which inhibits vascular smooth muscle cell proliferation, said Gax protein having a molecular weight of about 30 to 36kDa and comprising a homeodomain, wherein the amino acid sequence of said homeodomain comprises amino acid 185 through amino acid 245 of the amino acid sequence set forth in SEQ. ID. NO. 4.

Please cancel claims 2-4, 6-12, 14-17, 19-22, 24-27 and 29-32 without prejudice or disclaimer.

Please add the following claims:

33. (New) An isolated nucleic acid molecule comprising a nucleotide sequence encoding a Gax protein that inhibits vascular smooth muscle cell proliferation, said Gax protein having an amino acid sequence consisting essentially of the amino acid sequence set forth in SEQ. ID. NO. 2 or the amino acid sequence set forth in SEQ. ID. NO. 4.
34. (New) An isolated nucleic acid molecule encoding a Gax protein that inhibits vascular smooth muscle cell proliferation, said nucleic acid molecule comprising a nucleotide sequence encoding an amino acid sequence selected from the group consisting of the following:
  - (a) amino acid 223 through amino acid 302 of the sequence set forth in SEQ. ID. NO. 4;
  - (b) amino acid 67 through amino acid 222 of the sequence set forth in SEQ. ID. NO. 4;
  - (c) amino acid 30 through amino acid 66 of the sequence set forth in SEQ. ID. NO. 4;
  - (d) amino acid 1 through amino acid 29 of the sequence set forth in SEQ. ID. NO. 4; and combinations thereof.
35. (New) The isolated nucleic acid molecule of claim 34 wherein said Gax protein has a molecular weight of from about 30 kDa to about 36 kDa.

36. (New) A process for preparing an isolated nucleic acid molecule encoding a Gax protein that inhibits vascular smooth muscle cell proliferation comprising the following steps:

- (a) providing a human cDNA library which comprises a cDNA molecule encoding a Gax protein;
- (b) amplifying said cDNA molecule by a polymerase chain reaction using said cDNA molecule as a template and a pair of primers which bind under stringent conditions to the 5' end and the 3' end of the open reading frame sequence of SEQ ID. NO. 4, and
- (c) isolating the product of the polymerase chain reaction.

37. (New) An isolated DNA molecule prepared by the process of claim 45.

38. (New) An isolated DNA molecule with a Gax protein comprising an amino acid sequence which is at least 97% identical to the amino acid sequence set forth in SEQ. ID. NO. 2 or the amino acid sequence set forth in SEQ. ID. NO. 4.

39. (New) The DNA molecule of claim 47 wherein said Gax protein is a rat Gax protein and has a molecular weight of from about 30 kDa to about 36 kDa.

40. (New) The DNA molecule of claim 47 wherein said Gax protein is a human Gax protein and has a molecular weight of from about 30 kDa to about 36 kDa.

#### REMARKS

Claims 1-32 are pending in this application. Claims 2-27 are hereby canceled. The specification and claim 1 is amended. Attached hereto is a document entitled "VERSION WITH MARKINGS TO SHOW CHANGES MADE" showing the additions as underlined and the deletions in brackets is attached hereto. New claims 33 to 40 are hereby added.

The specification has been amended to correct typographical errors and to add SEQ. ID. NOS. for the nucleotide sequences and amino acid sequences recited therein. The amendments add no new matter.

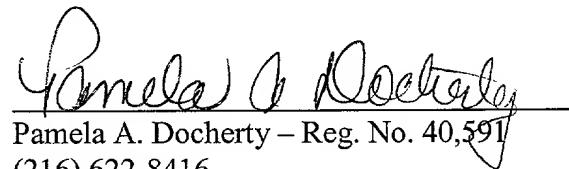
The substitute sequence listing which is attached hereto is the same as the computer readable form of the sequence listing which was filed on August 19, 1997, for the parent application. The substitute sequence listing adds no new matter.

Support for amended claim 1 and new claims 39 and 40 is found on page 9, line 10 and/or lines 18-20. Support for new claim 33 is found on page 4, lines 6-8 and 20-22, and page 32, lines 26-30. Support for new claims 34 and 35 is found on page 15, lines 5-17 and Figures 3 and 4. Support for new claims 36 and 37 is found on page 12, lines 30-37 through page 16, lines 1-15, and page 22, lines 17-26. Support for new claims 38-40 is found in Figs 1 and 3 and SEQ ID NO.s 2 and 4.

It is submitted that the amended claims and new claims add no new matter.

Respectfully submitted,

Date: August 27, 2001

  
\_\_\_\_\_  
Pamela A. Docherty – Reg. No. 40,591  
(216) 622-8416

VERSION WITH MARKINGS TO SHOW CHANGES MADE

1. (Amended) [A] An isolated DNA [sequence] encoding a mammalian Gax protein [or portion thereof] which inhibits vascular smooth muscle cell proliferation, said Gax protein having a molecular weight of about 30 to 36kDa and comprising a homeodomain, wherein the amino acid sequence of said homeodomain comprises amino acid 185 through amino acid 245 of the amino acid sequence set forth in SEQ. ID. NO. 4.
28. A method for inhibiting the proliferation of eukaryotic cells, comprising the following steps:
  - a. providing a nucleic acid sequence encoding a Gax protein, or portion thereof, which inhibits smooth muscle cell proliferation; and
  - b. administering said nucleic acid or said protein to the cells.

SUBSTITUTE  
SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Gorski, David H.  
Walsh, Kenneth

(ii) TITLE OF INVENTION: Growth Arrest Homeobox Gene

(iii) NUMBER OF SEQUENCES: 19

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: Calfee, Halter, and Griswold
- (B) STREET: 800 Superior Avenue
- (C) CITY: Cleveland
- (D) STATE: Ohio
- (E) COUNTRY: U.S.A.
- (F) ZIP: 44114-2688

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER:
- (B) FILING DATE:
- (C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Golrick, Mary E.
- (B) REGISTRATION NUMBER: 34829
- (C) REFERENCE/DOCKET NUMBER: 22311/00114

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (216) 622-8200
- (B) TELEFAX: (216) 241-0816
- (C) TELEX: 980499

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2244 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 197..1108

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTCAAGTGT TATACTGCA GGAGACTGGC CGCTCGGCTC AGGACTGGGA TTAGCGGGCT	60
CTGCTCAAAC CCGCGCGGCT TTTACATTAG GAGTGAGTGG GGGAGAGTCC TAGGATTCT	120
AGTAAAAAGT GACAGCGCTT GGTGGACTTT GGGACCTTCG TGAAGTCTTC TGCTTGGAAAG	180
CTGAGACTTG CATGCC ATG GAA CAC CCC CTC TTT GGC TGC CTG CGC AGC	229
Met Glu His Pro Leu Phe Gly Cys Leu Arg Ser	
1                      5                      10	
CCC CAC GCC ACA GCG CAA GGC TTG CAC CCC TTC TCG CAG TCT TCT CTG	277
Pro His Ala Thr Ala Gln Gly Leu His Pro Phe Ser Gln Ser Ser Leu	
15                    20                    25	
GCC CTC CAT GGA AGA TCT GAC CAC ATG TCC TAC CCC GAA CTC TCC ACA	325
Ala Leu His Gly Arg Ser Asp His Met Ser Tyr Pro Glu Leu Ser Thr	
30                    35                    40	
TCT TCC TCG TCT TGC ATA ATC GCG GGA TAC CCC AAT GAG GAG GGC ATG	373
Ser Ser Ser Cys Ile Ile Ala Gly Tyr Pro Asn Glu Glu Gly Met	
45                    50                    55	
TTT GCC AGC CAG CAT CAC AGG GGG CAC CAC CAC CAC CAC CAC CAC	421
Phe Ala Ser Gln His His Arg Gly His His His His His His His	
60                    65                    70                    75	
CAT CAC CAC CAC CAG CAG CAG CAC CAG GCT CTG CAA AGC AAC TGG	469
His His His Gln Gln Gln His Gln Ala Leu Gln Ser Asn Trp	
80                    85                    90	
CAC CTC CCG CAG ATG TCC TCC CCG CCA AGC GCG GCC CGG CAC AGC CTT	517
His Leu Pro Gln Met Ser Ser Pro Pro Ser Ala Ala Arg His Ser Leu	
95                    100                    105	
TGC CTG CAG CCT GAT TCC GGA GGG CCC CCG GAG CTG GGG AGC AGC CCT	565
Cys Leu Gln Pro Asp Ser Gly Gly Pro Pro Glu Leu Gly Ser Ser Pro	
110                    115                    120	
CCG GTC CTG TGC TCC AAC TCT TCT AGC CTG GGC TCC AGC ACC CCG ACC	613
Pro Val Leu Cys Ser Asn Ser Ser Leu Gly Ser Ser Thr Pro Thr	
125                    130                    135	
GGA GCC GCG TGC GCA CCA AGG GAT TAT GGC CGT CAA GCG CTG TCA CCC	661
Gly Ala Ala Cys Ala Pro Arg Asp Tyr Gly Arg Gln Ala Leu Ser Pro	

PDB ID: 3D9B

140	145	150	155		
GCA GAA GTG GAG AAG AGA AGT GGC AGC AAA AGA AAA AGC GAC AGT TCA Ala Glu Val Glu Lys Arg Ser Gly Ser Lys Arg Lys Ser Asp Ser Ser	160	165	170	709	
GAT TCC CAG GAA GGA AAT TAC AAG TCA GAA GTG AAC AGC AAA CCT AGG Asp Ser Gln Glu Gly Asn Tyr Lys Ser Glu Val Asn Ser Lys Pro Arg	175	180	185	757	
AGG GAA AGA ACA GCT TTC ACC AAA GAG CAA ATC AGA GAA CTT GAG GCA Arg Glu Arg Thr Ala Phe Thr Lys Glu Gln Ile Arg Glu Leu Glu Ala	190	195	200	805	
GAG TTC GCC CAT CAT AAC TAT CTG ACC AGA CTG AGA AGA TAT GAG ATA Glu Phe Ala His His Asn Tyr Leu Thr Arg Leu Arg Arg Tyr Glu Ile	205	210	215	853	
GCG GTG AAC CTA GAC CTC ACT GAA AGA CAG GTG AAA GTG TGG TTC CAG Ala Val Asn Leu Asp Leu Thr Glu Arg Gln Val Lys Val Trp Phe Gln	220	225	230	235	901
AAC AGG AGA ATG AAG TGG AAG CGG GTC AAG GGG GGA CAA CAA GGA GCT Asn Arg Arg Met Lys Trp Lys Arg Val Lys Gly Gly Gln Gln Gly Ala	240	245	250	949	
GCA GCC CGA GAA AAG GAA CTG GTG AAT GTG AAA AAG GGA ACA CTT CTT Ala Ala Arg Glu Lys Glu Leu Val Asn Val Lys Lys Gly Thr Leu Leu	255	260	265	997	
CCA TCA GAG CTG TCA GGA ATT GGT GCA GCC ACC CTC CAG CAG ACA GGG Pro Ser Glu Leu Ser Gly Ile Gly Ala Ala Thr Leu Gln Gln Thr Gly	270	275	280	1045	
GAC TCA CTA GCA AAT GAC GAC AGT CGC GAT AGT GAC CAC AGC TCT GAG Asp Ser Leu Ala Asn Asp Asp Ser Arg Asp Ser Asp His Ser Ser Glu	285	290	295	1093	
CAC GCA CAC TTA TGATACATAC AGAGACCAGC TCCGTTCTCA GGAAAGCACC His Ala His Leu	300			1145	
ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGGT				1205	
CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTCTC ATGATTGATA GAAGGTTTAC				1265	
ACTAAGTGCC TCTTATTGAA GATGCTTCCA CAGTGAAATT GGAGAAAGTG AACATATCTA				1325	
AATATACTTG TTCCTTATAT GACAGAGAGG GAGATGAATG TTTGCTTGG CTTGCACTGA				1385	
AAATTAAATT GCTACCAAGA GCAAACTCGG TAAGACATTT TGACTCAAGT TGTCTCCAGA				1445	
GTGAAGATGT TATAGAAATG CTTTGAACAT TCCAGTTGTA CCAGGTCATG TGTGTGACAC				1505	

TGGGCAGGTA TTTGCTTTG CTTGCACTGA AACTTAAACT GCTATCAAGT TAACCCATGA 1565  
 AATAGTTAT CTTGAACAGC CACAGTGCCT GAAATCACCA AGTGGATATA AAATGAAC TG 1625  
 AAATTCTGT A TATATTACTC CTAAGTCATT TTCCTGTCTT CACTAATT AGCAAATGCA 1685  
 TTCATATTAG CTGATGAAAA TAGGCTTCC CGTGGACAAA TGCAGCCAGC TTCTTGTATT 1745  
 TTTATACATT TTTTGTCA G TCAGAGACAT CAGTATGTGC TTACTTGTGT TCAAGTAGAG 1805  
 GAAATGCAGT AGAGTCTGAT AGGACATATT CTTGGTACCA CAGACAAA AC AAATCTTCTG 1865  
 TTGCATTGAC TATCAACTGC TGCAGATACA TTAGAGAAC A CACCTAGCCC CCCTCCAGCC 1925  
 TCCCTCTGTT ATCGCTCGAA GACATTAGCG TCATAGGCAA GTAGTTACCT TGCCAAATGA 1985  
 GTCTTGTGTG GCAGATGTCT GATTTGTAT CTTTAAACTG TTAATGGTAT GTGTCTGCTT 2045  
 CAGTTAACAG GGAAAAAGAT TTCTTCCTCA TTGTTTATGA TACAAAACCC AAGTGCCAAA 2105  
 CAAAGCTAGT TCTTCAAGGG ATAGATGAGA AACTGAATGT CTGACAAGTA GACTCAGCGA 2165  
 AAATACATTA TTTTCAGAG GCTGTGTATT CATGCAGTAC AAGTCCTTGT ATTTGTAAA 2225  
 AAAAAAAAGTT AAATAAATG 2244

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 303 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Glu	His	Pro	Leu	Phe	Gly	Cys	Leu	Arg	Ser	Pro	His	Ala	Thr	Ala
1				5					10					15	
Gln	Gly	Leu	His	Pro	Phe	Ser	Gln	Ser	Ser	Leu	Ala	Leu	His	Gly	Arg
					20			25					30		
Ser	Asp	His	Met	Ser	Tyr	Pro	Glu	Leu	Ser	Thr	Ser	Ser	Ser	Cys	
			35			40							45		
Ile	Ile	Ala	Gly	Tyr	Pro	Asn	Glu	Glu	Gly	Met	Phe	Ala	Ser	Gln	His
					50			55					60		
His	Arg	Gly	His	Gln											
					65			70					75		80
Gln	Gln	Gln	His	Gln	Ala	Leu	Gln	Ser	Asn	Trp	His	Leu	Pro	Gln	Met

85	90	95
Ser Ser Pro Pro Ser Ala Ala Arg His Ser Leu Cys Leu Gln Pro Asp		
100	105	110
Ser Gly Gly Pro Pro Glu Leu Gly Ser Ser Pro Pro Val Leu Cys Ser		
115	120	125
Asn Ser Ser Ser Leu Gly Ser Ser Thr Pro Thr Gly Ala Ala Cys Ala		
130	135	140
Pro Arg Asp Tyr Gly Arg Gln Ala Leu Ser Pro Ala Glu Val Glu Lys		
145	150	155
Arg Ser Gly Ser Lys Arg Lys Ser Asp Ser Ser Asp Ser Gln Glu Gly		
165	170	175
Asn Tyr Lys Ser Glu Val Asn Ser Lys Pro Arg Arg Glu Arg Thr Ala		
180	185	190
Phe Thr Lys Glu Gln Ile Arg Glu Leu Glu Ala Glu Phe Ala His His		
195	200	205
Asn Tyr Leu Thr Arg Leu Arg Arg Tyr Glu Ile Ala Val Asn Leu Asp		
210	215	220
Leu Thr Glu Arg Gln Val Lys Val Trp Phe Gln Asn Arg Arg Met Lys		
225	230	235
Trp Lys Arg Val Lys Gly Gly Gln Gln Gly Ala Ala Ala Arg Glu Lys		
245	250	255
Glu Leu Val Asn Val Lys Lys Gly Thr Leu Leu Pro Ser Glu Leu Ser		
260	265	270
Gly Ile Gly Ala Ala Thr Leu Gln Gln Thr Gly Asp Ser Leu Ala Asn		
275	280	285
Asp Asp Ser Arg Asp Ser Asp His Ser Ser Glu His Ala His Leu		
290	295	300

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 941 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: both
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 33..941

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GTCTTCTACC TGGAACCGA AACTTGCATG CT ATG GAA CAC CCG CTC TTT GGC Met Glu His Pro Leu Phe Gly	53
1 5	
TGC CTG CGC AGC CCT CAC GCC ACG GCG CAA GGC TTG CAC CCG TTC TCC Cys Leu Arg Ser Pro His Ala Thr Ala Gln Gly Leu His Pro Phe Ser	101
10 15 20	
CAA TCC TCT CTC GCC CTC CAT GGA AGA TCT GAC CAT ATG TCT TAC CCC Gln Ser Ser Leu Ala Leu His Gly Arg Ser Asp His Met Ser Tyr Pro	149
25 30 35	
GAG CTC TCT ACT TCT TCC TCA TCT TGC ATA ATC GCG GGA TAC CCC AAC Glu Leu Ser Thr Ser Ser Cys Ile Ile Ala Gly Tyr Pro Asn	197
40 45 50 55	
GAA GAG GAC ATG TTT GCC AGC CAG CAT CAC AGG GGG CAC CAC CAC CAC Glu Glu Asp Met Phe Ala Ser Gln His His Arg Gly His His His His	245
60 65 70	
CAC CAC CAC CAT CAC CAC CAT CAG CAG CAG CAG CAC CAG GCT CTG CAA His His His His His Gln Gln Gln His Gln Ala Leu Gln	293
75 80 85	
ACC AAC TGG CAC CTC CCG CAG ATG TCT TCC CCA CCG AGT GCG GCT CGG Thr Asn Trp His Leu Pro Gln Met Ser Ser Pro Pro Ser Ala Ala Arg	341
, 90 95 100	
CAT AGC CTC TGC CTC CAG CCC GAC TCT GGA GGG CCC CCA GAG TTG GGG His Ser Leu Cys Leu Gln Pro Asp Ser Gly Gly Pro Pro Glu Leu Gly	389
105 110 115	
AGC AGC CCG CCC GTC CTG TGC TCC AAC TCT TCC AGC TTG GGC TCC AGC Ser Ser Pro Pro Val Leu Cys Ser Asn Ser Ser Leu Gly Ser Ser	437
120 125 130 135	
ACC CCG ACT GGG GCC GCG TGC GCG CCG GGG GAC TAC GGC CGC CAG GCA Thr Pro Thr Gly Ala Ala Cys Ala Pro Gly Asp Tyr Gly Arg Gln Ala	485
140 145 150	
CTG TCA CCT GCG GAG GCG GAG AAG CGA AGC GGC GGC AAG AGG AAA AGC Leu Ser Pro Ala Glu Ala Glu Lys Arg Ser Gly Gly Lys Arg Lys Ser	533
155 160 165	

GAC AGC TCA GAC TCC CAG GAA GGA AAT TAC AAG TCA GAA GTC AAC AGC 581  
 Asp Ser Ser Asp Ser Gln Glu Gly Asn Tyr Lys Ser Glu Val Asn Ser  
 170 175 180

AAA CCC AGG AAA GAA AGG ACA GCA TTT ACC AAA GAG CAA ATC AGA GAA 629  
 Lys Pro Arg Lys Glu Arg Thr Ala Phe Thr Lys Glu Gln Ile Arg Glu  
 185 190 195

CTT GAA GCA GAA TTT GCC CAT CAT AAT TAT CTC ACC AGA CTG AGG CGA 677  
 Leu Glu Ala Glu Phe Ala His His Asn Tyr Leu Thr Arg Leu Arg Arg  
 200 205 210 215

TAC GAG ATA GCA GTG AAT CTG GAT CTC ACT GAA AGA CAG GTA AAA GTC 725  
 Tyr Glu Ile Ala Val Asn Leu Asp Leu Thr Glu Arg Gln Val Lys Val  
 220 225 230

TGG TTC CAA AAC AGG CGG ATG AAG TGG AAG AGG GTA AAG GGT GGA CAG 773  
 Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Arg Val Lys Gly Gly Gln  
 235 240 245

CAA GGA GCT GCG GCT CGG GAA AAG GAA CTG GTG AAT GTG AAA AAG GGA 821  
 Gln Gly Ala Ala Ala Arg Glu Lys Glu Leu Val Asn Val Lys Lys Gly  
 250 255 260

ACA CTT CTC CCA TCA GAG CTG TCG GGA ATT GGT GCA GCC ACC CTC CAG 869  
 Thr Leu Leu Pro Ser Glu Leu Ser Gly Ile Gly Ala Ala Thr Leu Gln  
 265 270 275

CAA ACA GGG GAC TCT ATA GCA AAT GAA GAC AGT CAC GAC AGT GAC CAC 917  
 Gln Thr Gly Asp Ser Ile Ala Asn Glu Asp Ser His Asp Ser Asp His  
 280 285 290 295

AGC TCA GAG CAC GCC CAC CTC TGA 941  
 Ser Ser Glu His Ala His Leu  
 300

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 302 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Glu His Pro Leu Phe Gly Cys Leu Arg Ser Pro His Ala Thr Ala  
 1 5 10 15

Gln Gly Leu His Pro Phe Ser Gln Ser Ser Leu Ala Leu His Gly Arg  
 20 25 30

Ser Asp His Met Ser Tyr Pro Glu Leu Ser Thr Ser Ser Ser Cys  
     35                        40                        45

Ile Ile Ala Gly Tyr Pro Asn Glu Glu Asp Met Phe Ala Ser Gln His  
     50                        55                        60

His Arg Gly His Gln Gln  
     65                        70                        75                        80

Gln Gln His Gln Ala Leu Gln Thr Asn Trp His Leu Pro Gln Met Ser  
     85                        90                        95

Ser Pro Pro Ser Ala Ala Arg His Ser Leu Cys Leu Gln Pro Asp Ser  
     100                       105                       110

Gly Gly Pro Pro Glu Leu Gly Ser Ser Pro Pro Val Leu Cys Ser Asn  
     115                       120                       125

Ser Ser Ser Leu Gly Ser Ser Thr Pro Thr Gly Ala Ala Cys Ala Pro  
     130                       135                       140

Gly Asp Tyr Gly Arg Gln Ala Leu Ser Pro Ala Glu Ala Glu Lys Arg  
     145                       150                       155                       160

Ser Gly Gly Lys Arg Lys Ser Asp Ser Ser Asp Ser Gln Glu Gly Asn  
     165                       170                       175

Tyr Lys Ser Glu Val Asn Ser Lys Pro Arg Lys Glu Arg Thr Ala Phe  
     180                       185                       190

Thr Lys Glu Gln Ile Arg Glu Leu Glu Ala Glu Phe Ala His His Asn  
     195                       200                       205

Tyr Leu Thr Arg Leu Arg Arg Tyr Glu Ile Ala Val Asn Leu Asp Leu  
     210                       215                       220

Thr Glu Arg Gln Val Lys Val Trp Phe Gln Asn Arg Arg Met Lys Trp  
     225                       230                       235                       240

Lys Arg Val Lys Gly Gly Gln Gln Gly Ala Ala Ala Arg Glu Lys Glu  
     245                       250                       255

Leu Val Asn Val Lys Lys Gly Thr Leu Leu Pro Ser Glu Leu Ser Gly  
     260                       265                       270

Ile Gly Ala Ala Thr Leu Gln Gln Thr Gly Asp Ser Ile Ala Asn Glu  
     275                       280                       285

Asp Ser His Asp Ser Asp His Ser Ser Glu His Ala His Leu  
     290                       295                       300

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: modified\_base  
(B) LOCATION: 6  
(D) OTHER INFORMATION: /mod\_base= i

(ix) FEATURE:

- (A) NAME/KEY: modified\_base  
(B) LOCATION: 21  
(D) OTHER INFORMATION: /mod\_base= i

(ix) FEATURE:

- (A) NAME/KEY: modified\_base  
(B) LOCATION: 24  
(D) OTHER INFORMATION: /mod\_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AARATWTGGT TYCARAAYMG WMGWATGAA

29

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(ix) FEATURE:

- (A) NAME/KEY: modified\_base  
(B) LOCATION: 4  
(D) OTHER INFORMATION: /mod\_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TCAWARRTGW GCRTGYTC

18

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 30 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GCGCGCAGAT CTCACTGAAA GACAGGTAAA

30

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TTTACCTGTC TTTCAGTGAG

20

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 32 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCGCGCAGAT CTAGATTACAC TGCTATCTCG TA

32

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GCGCGTGCC CCGCTGATGC TGGCTGGCAA ACATGT

36

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GCGCGCTCTT GAAGGGCGAG AGAGGATTGG GA

32

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CTGGTTCGGC CCACCTCTGA AGGTTCCAGA ATCGATAG

38

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GGAGACTTCC AAGGTCTTAG CTATCACTTA AGCAC

35

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCGCGCGTCG ACGAACACCC CCTCTTGCG

30

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCGCGCAAGC TTTCATAGT GTGCGTGCTC

30

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CCCGCGCGGC TTTTACATTA GGAGT

25

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

Ö (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GCTGGCAAAC ATGCCCTCCT CATTG

25

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TGATGGCATG GACTGTGGTC ATGA

24

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TGATGGCATG GACTGTGGTC ATGA

24